Discrimination of coastal land cover features in polarimetric SAR imagery

Randy John N. Vinluan, Enrico C. Paringit and Epifanio D. Lopez

Department of Geodetic Engineering, College of Engineering
University of the Philippines, Diliman, Quezon City, The Philippines
Fax: (63 2) 434-3641; E-mail: rjnv@engg.upd.edu.ph

1. Introduction

The detection of land cover features is one of the earliest applications of imaging radar since its earliest deployment for civilian applications. Imaging radar has the distinct advantage of all-weather performance and day and night data acquisition capability over other remote sensing instruments operating in the visible and infrared portion of the electromagnetic spectrum. For this reason, imaging radar is extremely useful in tropical countries which are perennially covered by clouds like the Philippines.

In one of the earliest applications of radar imagery for natural resource assessment, it was reported that radar provides a means for delimiting varying associations of physical and cultural phenomena through the outlining image variations in tone, texture, pattern and shape. Also, image patterns delimited on radar were visually correlated with known, observable variations of physical and cultural phenomena. It was demonstrated early on that some types of crops are easily distinguishable based on the strength of the radar return signal. Later on, in order to improve classification accuracy, extensions of the classification of single date of radar imagery at one frequency and polarization have been done. These include the analysis of multi-date radar imagery, integration with optical data and the inclusion of derived parameters such as textural and fractal measures in the analysis.

The advent of polarimetric synthetic aperture radar (SAR) presented an opportunity to improve land cover classification results significantly through simultaneous observations on ground targets using three radar frequencies and four polarizations. In late 1996, the National Aeronautics and Space Administration of the United States deployed the airborne SAR (AIRSAR) instrument over several countries in the Pacific Rim, including the Philippines.

In this paper, we examine the performance of two multivariate statistical methods - discriminant analysis and principal components analysis - in the identification of several land cover features present in a coastal area. Specifically, we first determine which radar bands would make up the land cover discrimination model and how much each band contributes to the model. Then we also try to identify the principal axes of variability within the data and determine which bands contribute significantly to the image variance. Finally, we compare the classification performance of the generated discriminant and principal components models with that of the conventional maximum likelihood classification.

The AIRSAR system imaged the project site located in the southeastern portion of Western Visayas region about 250 miles south-southwest of Manila on 26 November 1996. For this study, a subset of 8 km by 7 km was used. The study area is composed of extensive mudflats most of which have been structured into fishponds. Inland areas are covered by vast tracts of cultivated zones planted mainly with rice and lined up by settlements on the verge of urbanization. Brisk economic activities, mainly rooted in aquaculture, take place in this area owing to its strategic location, with the Guimaras-Iloilo Strait serving as a busy sea thoroughfare for vessels plying domestic routes. The capital city of Iloilo and its neighboring municipalities experience rapid development and consequently bring stress to its fragile coastal environment, thus requiring better methods to monitor its changes, including land cover. The study assumes reciprocity, that is HV=VH, and thus, only nine of the 12 available bands are used in the analysis.

2. Methods

Several techniques to analyze polarimetric SAR data for earth science applications have already been developed based on several radar parameters and characteristics of the radar instruments. This study makes use of multivariate statistical methods for the processing of backscatter values obtained using various radar wavelengths and polarizations. The first is called discriminant analysis. It is performed in order to describe the inherent separation existing between two or more variables, in this case, land cover types, based on a set of observations. This descriptive component of discriminant analysis is accomplished by computing linear functions, or discriminant functions, that best separate a set of variables. The discriminant functions represent coordinate axes in the p-dimensional space defined by p bands making up the data. The relationship between the p bands and the discriminant function axes is derived and the coordinates of the individual pixel vectors computed in terms of the discriminant function axes. Then, a set of classification functions, one for each class, are derived. A case is classified to the class associated with the classification function which returns the highest value. The discriminant analysis technique is a data reduction technique in that it seeks to find an optimal combination of variables that would maximize the separation between groups. It assumes that the data follow a multivariate normal distribution and that the variance-covariance matrices are statistically equivalent. However, violations of these assumptions will not usually lead to fatal results.

Principal components analysis is used to compress multivariate data sets by calculating a new coordinate system oriented in the long dimension of the distribution and a second axis perpendicular to the first. In general terms, the variance-covariance matrix S of the bands is computed. The eigenvalues, which give the lengths of the principal axes of the ellipsoid whose shape is defined by S, are then computed. Associated with each eigenvalue is a set of coordinates, called eigenvectors, which define the direction of the associated principal axis. The eigenvectors and eigenvalues therefore describe the new coordinate system which is oriented in the direction of maximum variability. A maximum likelihood classifier is applied to the reduced scene which is composed of only the first four principal components in order to identify the coastal land cover types present in the scene.

Several training areas for each land cover type were defined on the image. Then three hundred pixels from these training areas for each land cover type were randomly selected. Of this number, half were used as training data while the other half was used to test the derived discrimination and principal components models.

3. Results

3.1 Some properties of the training samples

In general, for most land cover types, radar backscatter in the HH polarization is higher while radar backscatter in the VV polarization is lower in the C and L bands. However, the opposite is true in the P band. It is also apparent that polarimetry is important in the discrimination of vegetated areas and built-up areas in all bands. Mature rice displays high variability in radar backscatter in all bands. This is due to the fact that the mature rice are in the early ripening to near-harvesting stages. From the vegetative phase, radar backscatter reaches a maximum upon entering the early ripening phase and then decreases after that. Other land cover types, namely young rice, sparsely built-up and densely built-up areas, also exhibit high variability in the C and L bands.

3.2 Results of discriminant analysis

All nine radar bands contribute significantly to class separation. This may be due to the fact that the different polarizations and radar wavelengths are important in the identification of different land cover types, such as L-band for agricultural cover, P-band for soil cover and cross-polarized L- or P-band for forest cover. The variables PHH, CHH, CVV, LHH, CHV, PVV, PHV, LHV and LVV were entered in that order. The contribution of a variable to group separation may be inferred from the magnitude of the discriminant coefficients. The first function, which accounts for 70 per cent of group separation is dominated by CHH, PHH, CVV and CHV. The 12 a priori defined land cover classes appear distinctly. The second function, which accounts for 13 per cent of group separation, is dominated by PHV, CHV, PVV and CVV. It is comprised mostly of negative loadings resulting to a reversal of brightness intensity compared to the first. The third function, which accounts for 8 per cent of group separation, is dominated by PHH, LHH, LHV and PVV. The classification procedure yielded an overall accuracy rate of 84.56 per cent, or a kappa coefficient of 0.83.

3.3 Results of principal components analysis

The first three principal components account for 95 per cent of the total variance. The nine bands contribute positive values in the first component (PC1) where CHH, LVV and PHH already contribute 43 per cent to the component loading. A visual inspection of PC1, which explains 79 per cent of the variance, reveals the presence of the 12 land cover classes, like in the first method. PC2 account for 12 per cent of the variance and it is dominated by the same three variables that dominate PC1. Since PC2 has predominantly negative loadings, a reversal in pixel intensity from PC1 is observed. Bright lines and dark spots now appear on the water body. The bright lines may be due to disturbance due to aircraft movement. The image defined by the first four principal components was classified using the maximum likelihood algorithm. An overall accuracy rate of 81.72 per cent, or a kappa coefficient of 0.80 was achieved.

3.4 Comparison with maximum likelihood algorithm

The classification by maximum likelihood algorithm alone resulted to an overall accuracy rate of 84.78 per cent, or a kappa coefficient of 0.83. However, this difference in classification performance with discriminant analysis is not statistically significant. Further comparisons between the three classification techniques were made regarding the effect of (a) increasing number of bands in the classification, (b) increasing number of training samples per class, and (c) increasing number of identified land cover classes, on the accuracy of classification. Results showed that classification accuracy improves with increasing number of entered bands in the classification procedure. However, improvements beyond the first three bands (PHH, CHH and CVV) are not dramatic. Classification accuracy also improves with increasing number of training samples using the maximum likelihood algorithm and principal components analysis only. This implies that good results maybe achieved using classification by discriminant analysis using only a few training samples. Finally, classification accuracy decreases with increasing number of classes to be identified.

4. Conclusions

The study evaluated the performance of two multivariate statistical techniques - discriminant analysis and principal components analysis - in the identification of coastal land cover features in polarimetric SAR imagery. The results were then compared with that obtained using the conventional maximum likelihood algorithm. It was found that all bands in the polarimetric SAR data contributed significantly to the discriminant model. All bands dominated the first three discriminant functions. The first three discriminant functions explained 93 per cent of the discrimination between the 12 land cover classes identified. The 12 classes were evident in the image defined by the first discriminant function. The result of the classification procedure yielded an overall accuracy of 84.56 per cent, or a kappa coefficient of 0.83. The result of principal components analysis showed that the first three principal components accounted for 95 per cent of the image variance. Like in the first method, the 12 classes are evident in the first principal component. The result of the classification procedure yielded an overall accuracy of 81.72 per cent, or a kappa coefficient of 0.80. The classification by maximum likelihood algorithm resulted to an overall accuracy of 84.78 per cent, or a kappa coefficient of 0.83. While that seemed to be a slight improvement over the other two methods, there was actually no significant difference in that result compared with that obtained using discriminant analysis. However, if one's particular SAR application involved more than classification, the image transforms would be more useful since more information could be derived from those methods which could improve the classification or help in understanding the phenomena being studied.